PRE-SYNOPSIS OF THE PROPOSED DOCTORAL RESEARCH WORK

ON

An empirical study of proficient optimization Technique for Training Deep Neural Networks

For the Award of

Doctor of Philosophy in

Computer Science

Faculty of Science & Technology

Submitted By

Mr. Daimi Syed Asif Syed Asim

Under the Supervision of

Prof. Dr. Sudhir B. Jagtap

Principal & Research Supervisor

Swami Vivekanand Mahavidyalaya, Udgir

December 2024

CONTENT

Sr. No.	Content Details	Page No
1	Abstract	3
2	Introduction	3-4
3	Review of Literature	4
4	Statement of Research Problem	5
5	Research methodology	5
6	Importance & Significance of the Research	5
7	Optimization Algorithms	6
8	Conclusion	6
9	Tentative Chapter Scheme	7
	Chapter 1: Introduction	
	Chapter 2: Review of Literature	
	Chapter 3: Optimization Algorithms: An Overview	
	Chapter 4: Data Interpretation and Analysis	
	Chapter 5: Finding suggestion and Conclusions	
10	Bibliography/References	8-9
11	Working plan	10

1. Abstract

Deep learning is a sub-branch of artificial intelligence that acquires knowledge by training a neural network. It has many applications in the field of banking, automobile industry, agriculture, and healthcare industry. Deep learning has played a significant role in solving difficult tasks related to computer vision, such as image classification, natural language processing, and object detection. On the other hand, optimizers also play an intrinsic role in training the deep learning model. Recent studies have proposed many deep learning models, such as VGG, ResNet, DenseNet, and ImageNet. In addition, there are many optimizers such as stochastic gradient descent (SGD), Adam, AdaDelta, Adabelief, and AdaMax. For this study, i have to select those models that require lower hardware requirements and shorter training times, which facilitates the overall training process. i have to modify the Adam based optimizers and minimize the cyclic path.

I have removed an additional hyper-parameter from RMSProp and observed that the optimizer works with various models. The learning rate is set to minimum and constant. The initial weights are updated after each epoch, which helps to improve the accuracy of the model. We also changed the position of the epsilon in the default Adam optimizer. By changing the position of the epsilon, it accumulates the updating process. We used various models with SGD, Adam, RMSProp, and the proposed optimization technique. The results indicate that the proposed method is effective in achieving the accuracy and works well with the state-of-the-art architectures.

2. Introduction

Initially, the concept of neural networks arose independently. But with the pages of time, the emergence of new backgrounds and new disciplines has made the natural combination of neural network technology and other technologies form a set of computational processes with solid theories that are highly efficient in solving practical problems in various fields. Today neural networks are no longer considered a separate field of study but are part of new contexts and new disciplines [5]. Deep neural networks are standard tool for solving computer vision tasks. Deep neural networks are mostly used for image classification, natural language processing, and rarely for audio recognition. Image classification is one of the most fundamental tasks of computer vision. It has revolutionized and propelled technological advancements in the most prominent fields such as healthcare, banking, automobile industry, and agriculture. Image classification is the process of taking an input, e.g., a picture of a cat or

dog, and outputting a class of a picture. Image classification requires a huge dataset. There are many open-source datasets for image classification available online, such as CIFAR-10, CIFAR-100, MNIST, COCO, and ImageNet [6–8].

The main objective of any deep learning algorithm is to predict outputs closer to the actual output [9]. It helps to reduce the cost function, which is based on the prediction error. Optimization algorithms play an important role in the forward and backward propagation of a neural network during the training process. It speeds up the training process and finds the optimal parameters for a neural network. An optimization algorithm executes all possible solutions iteratively until it reaches a point that is optimum or satisfactory [10]. Optimal optimizers effectively help neural networks during the training process. Gradient descent is the most basic and popular optimization algorithm. Gradient descent is an optimization algorithm used to minimize some function by iteratively moving in the direction of the steepest descent.

3. Review of Literature:

The data which support the study will be from references from the educational websites: Google Scholar, Baidu Scholar, and cnki. This research is also be include some information and data provided by experts and professors in this area as a supplement.

There are various types of neural networks, such as recurrent neural network (RNN) and convolutional neural network. In recent years, deep learning has shown a great impact over the performance of the neural networks in many complicated learning tasks such as natural language processing, image classification, and object detection. A convolutional neural network is a class of artificial neural networks that is used to analyze visual imagery. CNNs have become the state-of-the-art computer vision technique. Recently, a lot of research has been conducted for optimizing parameters of deep neural network (DNNs). Gradient-based algorithms are widely used for this purpose. However, there have been issues with parameter optimization, such as vanishing gradients. The authors in Ref. [1] proposed a novel algorithm, named the evolved gradient direction optimizer (EVGO), to cope with the vanishing gradient. The goal of the researcher is to find a better search space and modify the gradient value by keeping it away from zero in order to maintain the learning smoothly.

4. Statement of the Research Problem

An empirical study of optimization Technique for Training Deep Neural Networks

5. Research Methodology:

The primary purpose of this study is to develop an optimization technique to improve the accuracy of neural networks. In addition, the optimization technique should reduce the time and complexity of the neural network. In this section, we will briefly explain the proposed methodology. there are three layers, i.e, the data pre-processing layer, the training and testing phase, and the evaluation of machine learning models. The data pre-processing layer deals with the dataset, the training and testing phase deals with the CNN models and optimizers, and the evaluation phase deals with the model accuracy.

6. Significance of the study:

Deep learning is important for the development of artificial intelligence because it allows machines to learn in the same way that humans do [17]. Traditional machine learning methods only rely on information about data that has been specifically programmed into the machine, meaning that it can take a lot of time and effort to train the machine. On the other hand, deep learning allows the machine to learn from its own experiences so that the machine can automatically adapt and learn new information on its own without requiring much human intervention. This capability can accelerate the development of artificial intelligence because it allows the machine to develop its own rules and strategies for problem solving without having to be explicitly programmed to do so. In the future, it may be possible for machines to independently develop their own intelligence without any human assistance at all [18].

7. Optimization Algorithms

Optimization plays an important role in the training of a neural network. Recent studies have proposed many optimization algorithms. Each optimization algorithm has its own advantages and disadvantages. In this section, we will briefly explain the current state-of-the-art optimizers. Optimization is an important process in the training of neural networks. The goal of optimization is to find a set of parameters that yields the best performance for a given problem. There are many different methods for performing optimization, each of which has advantages and disadvantages. The most common type of optimization used in neural networks is gradient descent. This method involves repeatedly adjusting the values of the network's parameters until the performance improves. Different types of problems can be optimized in different ways and by using different methods. In some cases, it is not possible to use gradient descent, which can make it necessary to use another method.

Training a neural network involves setting the parameters of its underlying model such that the network learns to perform the desired task. In order to train a network, it is necessary to define the loss function that will be used to evaluate the network's performance. A loss function is a function that takes the output of the network and assigns it a numerical value that represents its performance. The values of these parameters must be set so that the network can learn to perform the task well while minimizing the loss function value. The process of adjusting the values of the network's parameters to minimize the loss is known as the training procedure.

8. Conclusion

In this study, the aim to improve the optimization technique for training deep neural networks. Deep learning plays a major role in computer vision tasks such as image classification, natural language processing, and object detection. During deep neural network training, optimizers play an important role in improving the accuracy and reducing the training time. Optimizers are algorithms that are used to maximize the efficiency of the model or minimize the loss function. Recent studies have proposed many state-of-the-art optimizers to support the neural network training. In this study, I have to focus on different variants of optimization techniques and aim to introduce a optimization method to improve the accuracy of the deep neural networks.

Tentative Chapter Scheme

Academic research work is presented in several chapters and in this pattern following chapter scheme will be adopt by the researcher.

Chapter 1: Introduction

It will start with Introduction, research methodology and other necessary information will incorporate in this chapter.

Chapter 2: Review of Literature

Literature on related topics will be reviewed to understand the research area and such reviews will be arranged in this chapter.

Chapter 3: Optimization Algorithms: An Overview

This chapter will be analyzing the concept of artificial Intelligence Technology and its several applications.

Chapter 4: Data Interpretation and Analysis

This section explains the tools and technologies used in implementing and evaluating the proposed optimization technique during the training of neural networks. In this section, we will explain the experimental setup and results obtained during training and testing the deep neural network. In this section, we will discuss various performance metrics used during the training, validation, and testing phase of the different models. We will compare the results of each model and evaluate their performance.

Chapter 5: Finding suggestion and Conclusions

In this chapter an attempt will be made to enlist the findings of the study, suggestions and conclusion.

References:

1. Karabayir, I.; Akbilgic, O.; Tas, N. A novel learning algorithm to optimize deep neural networks: Evolved gradient direction optimizer (EVGO). IEEE Trans. Neural Netw. Learn. Syst. 2020, 32, 685–694.
2. Alzubaidi, L.; Zhang, J.; Humaidi, A.J.; Al-Dujaili, A.; Duan, Y.; Al-Shamma, O.; Santamaría, J.; Fadhel, M.A.; Al-Amidie, M.; Farhan, L. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. J. Big Data 2021, 8, 1–74. [CrossRef] [PubMed]
3. Wang, S.Y.; Wang, O.; Zhang, R.; Owens, A.; Efros, A.A. CNN-generated images are surprisingly easy to spot. . . for now. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 8695–8704.
4. Sharma, P.; Abrol, P. Analysis of Multiple Component Based CNN for Similar Citrus Species Classification. In Modern Approaches in Machine Learning & Cognitive Science: A Walkthrough; Springer: New York, NY, USA, 2022; pp. 221–232.
5. Prieto, A., Prieto, B., Ortigosa, E. M., Ros, E., Pelayo, F., Ortega, J., & Rojas, I. Neural networks: An overview of early research, current frameworks and new challenges. Neurocomputing, 214, 242-268 (2016).
6. Abouelnaga, Y.; Ali, O.S.; Rady, H.; Moustafa, M. Cifar-10: Knn-based ensemble of classifiers. In Proceedings of the 2016 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA, 15–17 December 2016; IEEE: New York, NY, USA, 2016; pp. 1192–1195.
7. Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Fei-Fei, L. Imagenet: A large-scale hierarchical image database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; IEEE: New York, NY, USA, 2009; pp. 248–255.
8. Lin, T.Y.; Maire, M.; Belongie, S.; Hays, J.; Perona, P.; Ramanan, D.; Dollár, P.; Zitnick, C.L. Microsoft coco: Common objects in context. In Proceedings of the European
9. Tian, C.; Fei, L.; Zheng, W.; Xu, Y.; Zuo, W.; Lin, C.W. Deep learning on image denoising: An overview. Neural Netw. 2020, 131, 251–275. [CrossRef] [PubMed]

10. Sun, R.Y. Optimization for deep learning: An overview. J. Oper. Res. Soc. China 2020, 8, 249–294. [CrossRef]
11. Zeiler, M.D. Adadelta: An adaptive learning rate method. arXiv 2012, arXiv:1212.5701.
12. Hinton, G. Neural networks for machine learning coursera video lecturesgeoffrey hinton. 2012. Available online: https: //archive.org/details/academictorrents_743c16a18756557a67478a7570baf24a59f9cda6 (accessed on 1 December 2022)
13. Sutskever, I.; Martens, J.; Dahl, G.; Hinton, G. On the importance of initialization and momentum in deep learning. In Proceedings of the International Conference on Machine Learning, PMLR, Atlanta, GA, USA, 17–19 June 2013; pp. 1139–1147.
14. Zhang, Z. Improved adam optimizer for deep neural networks. In Proceedings of the 2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS), Banff, AB, Canada, 4–6 June 2018; IEEE: New York, NY, USA, 2018; pp. 1–2.
15. Dozat, T. Incorporating nesterov momentum into adam. In Proceedings of the International Conference on Learning Representations, San Juan, Puerto Rico, 2–4 May 2016.
16. Barazandeh, B.; Curtis, K.; Sarkar, C.; Sriharsha, R.; Michailidis, G. On the convergence of adam-type algorithms for solving structured single node and decentralized min-max saddle point games. In Proceedings of the ICASSP 2022—2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Singapore, 22–27 May 2022; IEEE: New York, NY, USA, 2022; pp. 4343–4347
17. Jackson, P.C. Introduction to Artificial Intelligence; Courier Dover Publications: New York, NY, USA, 2019.
18. Haenlein, M.; Kaplan, A. A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. Calif. Manag. Rev. 2019, 61, 5–14. [CrossRef]
19. Pravitasari, A.A.; Iriawan, N.; Almuhayar, M.; Azmi, T.; Irhamah, I.; Fithriasari, K.; Purnami, S.W.; Ferriastuti, W. UNet-VGG16 with transfer learning for MRI-based brain tumor segmentation. TELKOMNIKA Telecommun. Comput. Electron. Control. 2020, 18, 1310–1318. [CrossRef]
20. Alom, M.Z.; Taha, T.M.; Yakopcic, C.; Westberg, S.; Sidike, P.; Nasrin, M.S.; Van Esesn, B.C.; Awwal, A.A.S.; Asari, V.K. The history began from alexnet: A comprehensive survey on deep learning approaches. arXiv 2018, arXiv:1803.01164

Working plan:

Year 1:

First 6 month's plan: Collection of previous studies and books on research topic and prepare the list of the respondents.

Second 6 month's plan: Reading related literature and start to write reviews.

Year 2:

First 6 month's plan: Complete the writing of review of related literature.

Second 6 month's plan: Complete the Data collection regarding research.

Year 3:

First 6 month's plan: To start the data analysis interpretation and write first chapter

Second 6 month's plan: Report writing printing, binding and submission of thesis

Research Student	Research Supervisor
Mr. Daimi Syed Asif Syed Asim	**Dr. Sudhir B. Jagtap**